Estimating the Effects of Wage Subsidies on the Labour Demand in West-Germany - A Conditional Difference-in-Differences Approach

Marco Caliendo*

This draft: March 16, 2001

Working Paper

Abstract Wage subsidies are used as one means to fight unemployment in Germany. Despite their growing importance evaluation studies are rather scarce and aim on the labour supply side. In this paper we estimate their effects on the labour demand using the IAB establishment panel. To account for the selectivity problem in microeconometric evaluation studies, we apply a conditional difference-in-differences approach, i.e. we use a matching approach to construct a suitable control group in a first step, followed by a difference-in-differences estimation. We concentrate on West German Establishments and the time period 1995 to 1999 and evaluate programmes including employment integration measures and job creation schemes. We do not find any positive long-run effects on the employment for participating establishments.

Keywords: Evaluation, Active Labour Market Policy, Wage Subsidies, Labour Demand, Matching, Difference-in-Differences
JEL Classification: C13, C23, C25, H43, J64, J68

Contents

1. Introduction

Germany has been plagued, as most OECD countries, by high and persistent unemployment since the early 1970s. Active labour market policies (ALMP) have been regarded as a suitable means of fighting this unacceptable situation, because they are capable of meeting efficiency and equity goals at the same time. The most important measures can be categorized in two broad groups: Training on the one hand and subsidized employment on the other. Regarding training measures there exist plenty evaluation studies, whereas the studies for wage subsidies are rather scarce. Furthermore nearly all of the latter studies focus on the effects of ALMP on the labour supply, using data sets as the German Socio Economic Panel, the East-German Labour Market Monitor or the Labour Market Monitor Sachsen Anhalt.[1] In this paper we will focus on the other side of the market, that is we try to estimate the effects of wage subsidies on the labour demand. The data set used is the IAB-establishment panel, a yearly survey of more than 4,000 establishments in West Germany. The panel was started in 1993 and since 1997 there have been 5,000 establishments from East Germany in the sample, too.

Every microeconometric evaluation of labour market programmes tries to answer the question if the interesting outcome variable for an individual unit is affected by participation in an ALMP programme or not. Relevant outcome variables on the labour supply side could be the future employment probability or the future earnings of an individual, whereas on the labour demand side we might look at the number of total employees in the firm. As most wage subsidies are targeted on persons with bad labour market prospects, we take into account heterogeneous labour and estimate the effects for different skill levels.

Basically we would like to know the difference between the value of the individual units outcome in the actual situation and its value if it had not participated in the programme. The fundamental evaluation problem arises because we never observe both states (participation and non-participation) for the same individual unit at the same time, i.e. one of the states is counterfactual. Therefore finding an adequate control group is necessary to make a comparison possible.

This is not easy because the participating units in programmes usually differ in more aspects than just participation from the non-participating units, i.e. the participation decision is selective. Simply taking the difference between their outcomes after participation will not reveal the true training impact, i.e. will lead to a biased estimate. This selection process might occur on observable or unobservable characteristics.

To solve the fundamental evaluation problem and construct an appropriate control group we use a conditional difference-in-differences approach. The first step of our approach comprises of a matching procedure to reduce the selection on observable characteristics, that is, we try to find individual units in the control group that are fairly similar to the units in the treatment group in all relevant characteristics. In a second step we apply a difference-in-differences estimator to extinguish all selection that occurred on unobservable characteristics.

The conditional difference-in-difference estimator is rather data demanding not only in the sense that we need information about the unit before and after the programme took place, but also because we need a large pool of non-participants to find suitable controls for our treatment group. The longer the period we examine the unit after the measure, the more we can say about long-term

---

[1] For a recent overview of these studies see Hagen and Steiner (2000) or Hujer and Caliendo (2000).

1

effects. Therefore we concentrate in a first approach on all establishments that participated in a measure in 1996 and have been surveyed from 1995-1999. Doing so, we can estimate the DiD-estimator for three years after the measure. As in 1995 only West-German establishments have been in the survey this study concentrates on West-Germany.

The analyzed measures are employment integration measures, employment assistance for the long-term unemployed and job creation schemes as well as structural adjustment schemes in the profit and non-profit sector. Unfortunately the number of observations for every measure has been too small, so that we had to pool some measures and could not take into account the problem of heterogeneous treatments.

The paper is organized as follows. The next section gives a brief overview of the existing wage subsidy programmes in Germany, before we discuss the theoretical effects of these programmes on the labour demand in section 3. Section 4 deals with previous empirical findings, whereas section 5 presents the IAB establishment panel and some descriptive results. Finally section 6 concentrates on the actual evaluation process, starting with defining our outcome variable of interest, before giving a short theoretical outline about the fundamental evaluation problem in microeconometric studies and the approach we apply to deal with it. We then turn to the estimation of the propensity score to participate in a wage subsidy programme and based on these results the matching procedure follows. The constructed matched sample will be used for our conditional difference-in-differences estimation. Section 7 concludes.

2. Wage Subsidies in Germany

Labour market policies in Germany are organized by the Federal Employment Office ('Bundesanstalt für Arbeit', FEO). Up to 1998 the legal basis for the labour market policy in Germany has been the work support act ('Arbeitsförderungsgesetz', AFG), founded in 1969. From there on, the new Social Code SGB III ('Sozialgesetzbuch') is playing this role. Changes have been made not only in the objectives, like a more intensive focus on problem groups of the labour market, but also in the institutional organization of labour market policy, leading to decentralization and more flexibility in the regional allocation of resources to different measures. As a result of these changes, the comparability of the different measures has suffered. A brief overview of AFG's historical evolution can be found in Staat (1997), whereas Fitzenberger and Speckesser (1998) and Hujer and Caliendo (2000) describe the most relevant reforms of the new SGB III.

We will now take a brief look on the spending on labour market policies in 1999. Table 1 shows that 44.8% of the total spending of the FEO have been dedicated to active measures, the relation in East Germany (56.9%) being much higher than in the West (37.1%). The most important measures have been the support of vocational training with 13.2 bn DM and subsidized employment, consisting of traditional job creation schemes ('Arbeitsbeschaffungsmaßnahmen') with 7.8 bn DM, structural adjustment schemes ('Strukturanpassungsmaßnahmen') with 5.05 bn DM and Employment Integration Measures ('Eingliederungszuschüsse') with 1.8 bn DM.

As vocational training aims directly to employees and therefore on the labour supply side it is of minor interest for our study.[2] Therefore we will focus on subsidized employment. The measures

---

[2] The interested reader should refer to Hujer, Maurer, and Wellner (1999) and Hujer and Wellner (2000) for an overview of vocational training under the AFG.

in this area can be categorized by their recipients, the targeted population, the size of the subsidy and their duration. An extensive overview of all these measures can be found in Hagen and Steiner (2000).

Table 1: **Spending on Labour Market Policies in 1999**

|	Germany		West		East	
	Spending		Spending		Spending	
	in bn DM	% of total	in bn DM	% of total	in bn DM	% of total
Total Spending Federal Employment Office	101.10	100.00	61.89	100.00	39.22	100.00
Active Labour Market Policies	45.30	44.81	22.98	37.13	22.32	56.91
Selected measures:						
Support of Vocational Training	13.20	13.06	7.77	12.55	5.43	13.84
Job Creation Schemes	7.81	7.73	2.14	3.46	5.66	14.43
Structural Adjustment Schemes (SAM)	1.48	1.46	0.25	0.40	1.23	3.14
SAM-East for Private Firms	3.57	3.53	0.14	0.23	3.43	8.75
Employment Integration Measures	1.84	1.82	1.39	2.25	0.46	1.17
Rehabilitation Measures	4.50	4.45	3.28	5.30	1.22	3.11
Free Support	1.09	1.08	0.50	0.81	0.59	1.50
Support of Professional Training	2.76	2.73	1.66	2.68	1.10	2.80
Crash Prog. against Youth Unempl.	1.90	1.88	1.13	1.83	0.78	1.99

Source: Bundesanstalt für Arbeit (2000)

Job creation schemes (JCS) are normally only available to non-profit organisations. They should support activities, which are of value for the society and additional in nature, that is without the subsidy they could not be executed. They include limited employment for long-term unemployed in projects, to improve their labour market prospects. Even though JCS should be co-financed measures, where between 30% and 75% of the costs are subsidies by the Federal Employment Office and the rest is paid by the implementing institution (public or private legal entities, mainly municipalities), exceptions can be made, in the direction of an higher subsidy-quota. The subsidy is normally paid for 12 months, but can be extended up to 24 months. Priority is given to projects, which improve the chances for permanent jobs that support structural improvement in social or environmental services or that aim at the integration of extremely hard-to-place individuals.

Especially in East Germany structural adjustment schemes (SAM) play a prominent role. Their goal is, analogous to JCS, the integration into regular employment, but less severe eligibility criteria apply to participants, so that individuals who are only threatened by unemployment may participate already. The SAM consist of a wage subsidy equal to the amount of unemployment allowance or assistance the individual would have received if unemployment had continued, typically paid for a maximum period of 36 months. In East Germany the SAM may be implemented by public institutions and private companies ("SAM Ost für Wirtschaftsunternehmen", SAM-East), whereas in West Germany only the first is possible.

Finally we will take a closer look on the wage subsidies directly paid to firms employing workers with poor labour market prospects. The employment integration measures (§§217-224 SGB III, EIM) summarize four formerly separated instruments of the AFG. The common goal is to help

Figure 1: **Entries into Different Wage Subsidy Programmes in West-Germany, 1991-99**
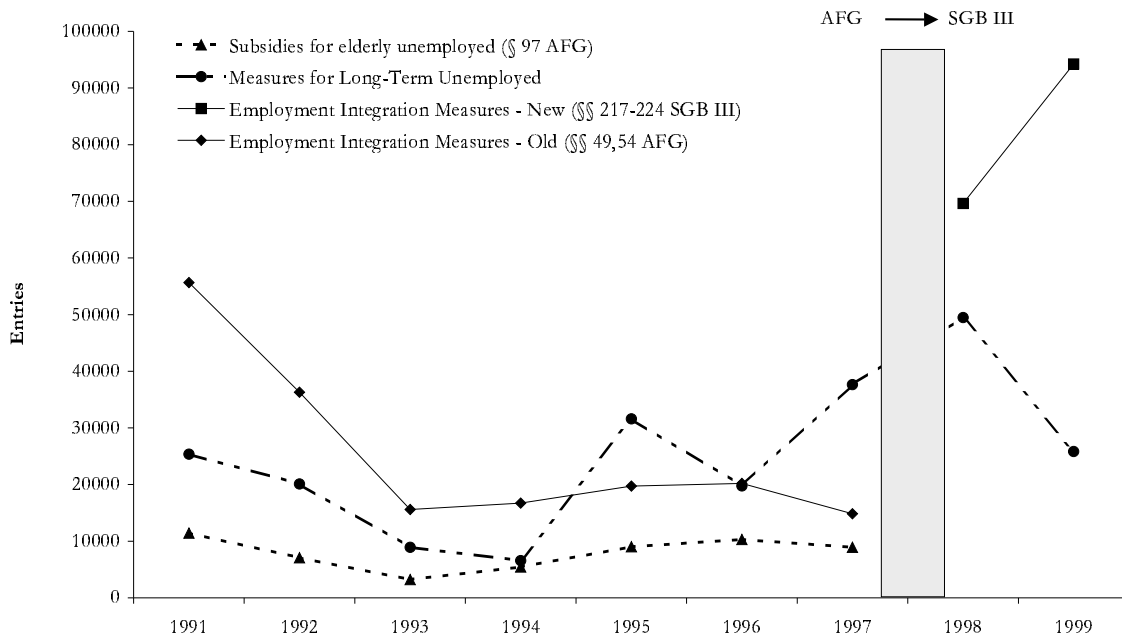
unemployed, familiarizing with a new job.[3] A good overview regarding this measures can be found in Buslei and Steiner (1999).

EIM comprise various types of wage subsidies. Settling-in grants can be paid, if the employee needs a special familiarizing after returning to the job. 30 % of the regular wage are normally paid as a subsidy for up to 6 months. There are also subsidies for hard-to-place unemployed people. In this case the subsidy amounts to 50 % of the wage and can be paid up to 12 months. Finally there are also measures for elderly (that is older than 55 years) and long-term unemployed, amounting to up to 50 % of the wage paid for up to 24 months. Rules have been set up to avoid substitution effects, e.g. the subsidy is not paid if the employment agency suspects, that the employer has set off another worker to employ the subsidized employee. Furthermore it is not allowed, that the employee has been employed in the last four years with the same employer. Another program worth mentioning is the employment assistance for the long-term unemployed. If a long-term unemployed is hired, the firm can claim a wage subsidy consisting of 60 to 80 % of the regular wage for up to one year.

Although the employment integration measures are relatively small compared to VT or JCS, they are very interesting for us, as they are directly paid to firms and therefore should have a direct impact on the labour demand of these firms. Figure 1 shows the number of entries into different kinds of wage subsidy programmes, not including JCS and SAM.[4]

In detail the entries into EIM, as well as subsidies for the long-term unemployed and elderly

---

[3]The instruments in the work support act have been: 'Einarbeitungszuschüsse' (§49 AFG), 'LKZ für ältere Arbeitslose' (§97 AFG), 'Eingliederungsbeihilfen' (§54 AFG) and 'Eingliederungshilfen' (§58 AFG).

[4]We do not include SAM and JCS at that point because they are not aimed at private firms and furthermore they are quantitative less important for our further study, as they are not requested very often. We do not have any request for SAM in our sample and only a few for JCS.

unemployed for West-Germany from 1991 to 1999 are depicted. The comparability suffers through the change from the AFG to SGB III (1997 → 1998). The entries into EIM under the AFG remained, after a heavy fall from 1991 to 1993, relatively constant between 15.000 and 20.000 entries per year from 1993 to 1997. The high number of entries in 1991 can be explained with special regulations for employees out of East Germany. To avoid substitution effects the regulations have been tightened leading to a decrease in the entries. With the introduction of the SGB III some new measures were summarized under EIM and the entries raised to nearly 100.000 in 1999. This expresses also a growing importance of these measures, the importance being higher in West Germany, because in East Germany there are also the SAM-East for private firms paying subsidies directly to private employers.

If we look at the entries into subsidies for elderly unemployed we can see, that the entries have been on a very low level in 1993, caused by financial restrictions of the FEO. From there on the entries grew to 10.000 a year. The same financial restrictions caused the drop of entries in measures for the long-term unemployed in 1993 and 1994. After that period the entries raised showing a growing importance of these measures, reaching a top in 1998 with nearly 50.000 entries a year.

3. Theoretical Effects of Wage Subsidies on the Labour Demand

In the following we will consider the theoretical effects of wage subsidies on the labour demand.[5] Thereby we will assume a simple static model in which a representative establishment produces one output y using two variable input factors, namely low- and high-skilled workers L_H and L_L:

$$y = F(L_H, L_L). \tag{1}$$

As usual it is assumed that $F_i > 0$, $F_{ii} < 0$ and $F_{ij} > 0$ with $i, j = L_L, L_H$. Additionally we assume that the two input factors considered are separable from other factors as e.g. capital, and that the labor supply is infinitely elastic. The optimal demand for high- and low- skilled workers is then set according to (point A in figure 1):

$$\frac{F_{L_H}}{F_{L_L}} = \frac{w_H}{w_L}. \tag{2}$$

Wage subsidies which are provided e.g. for the low skilled employees reduce the costs of employing these subsidized workers and will therefore have a positive effect on the labour demand.[6] In figure 1 the wage subsidies will rotate the isocost line from IC_1 to IC_2. If we assume that a proportional wage subsidy at a level z is paid on the wage of the low-skilled workers, the direct demand effect for this group is given by:
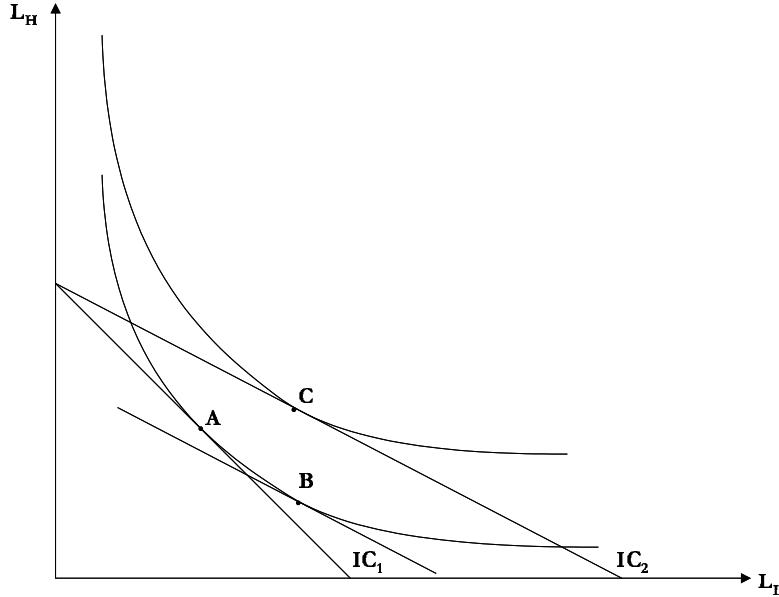
$$\Delta L_L = \{(1 - s)\sigma\} \cdot z > 0, \tag{3}$$

where ΔL_L reflects the increasing demand for low-skilled workers, s is the share of labor costs for low-skilled workers in the total production and σ measures the elasticity of substitution between high- and low-skilled workers.

---

[5] This section draws heavily on Hamermesh (1993).

[6] Steiner (2000), pp. 284-287.

Figure 2: **Theoretical Effects of Wage Subsidies**

At the same time the relative price for the high-skilled workers increases, therefore reducing the demand for this group of employees by:

$$\Delta L_H = -\{(1-s)\sigma\} \cdot z < 0. \tag{4}$$

However, with lower production costs and therefore decreasing prices the demand for the produced good will also increase, so that besides the direct substitution effect in (3) and (4) (point B in figure 1) one has also to take into account the scale effect. Thus the total demand effect for the two employment groups can be summarized as follows (point C in figure 1):

$$\Delta L_L = \{(1-s)\sigma + s\eta\} \cdot z > 0, \tag{5}$$

$$\Delta L_H = -\{(1-s)\sigma + (1-s)\eta\} \cdot z \begin{cases} > & 0 \\ = & 0, \\ < & 0 \end{cases} \tag{6}$$

where η measures the own price elasticity of demand. One can see that theory predicts that wage subsidies targeted at low-skilled workers will increase the demand for this group of employees whereas the effect on high-skilled workers is ambiguous and depends on whether the substitution or the scale effect prevails.

Since our framework is a static one, these theoretical considerations only hold in the long-run. They reflect the optimal response of an establishment to a change in the relative wage due to a wage subsidy. If we assume that there are costs associated with the adjustment of the labour demand, that these adjustment costs might differ between different skill groups and that

6

additionally the demand for high- and low-skilled workers are interdependently related, things get even more complicated and ambiguous.[7]

One might think e.g. that the adjustment costs for low-skilled workers are lower than for high-skilled. The demand for this group of workers would then adjust quicker to the changing relative prices whereas the response might be slower for the group of high-skilled workers.

Another problem arises if the implicitly made assumption that the wage subsidy covers the whole group of low-skilled workers is not fulfilled. If wage subsidy programmes are targeted only at a subgroup of the low-skilled, e.g. at the long-term unemployed, and if this subgroup is a substitute for other low-skilled, then displacement effects are likely to occur. In this case even the labour demand effect for the low-skilled becomes ambiguous so that an answer to these questions can only be given by an empirical investigation.

4. Previous Empirical Findings

Empirical studies which try to evaluate the employment effects of wage subsidies are rather scarce. In the following we will distinguish between studies evaluating employment effects on the labour supply and on the labour demand.[8]

The existing microeconometric studies which focus on the labour supply all use datasets for East Germany, namely the labour market monitors for East Germany and for Sachsen-Anhalt. Kraus and Steiner (1995) focus on the employment effects of JCS using the first dataset for the years 1990 to 1992. The authors conclude that in the long-run JCS have no effect regarding employment probabilities whereas in the short-run the effects of JCS are even negative. This result is confirmed by Hübler (1997) who used the same dataset for a period from 1990 to 1994 and by Kraus, Puhani, and Steiner (1999) who find negative effects of JCS for a period from 1992 to 1994, too.

The only study which finds positive effects, using a different methodological approach and the labour market monitor for Sachsen-Anhalt from 1991-1997, is the one by Eichler and Lechner (1999). Therefore it might be stated that there is some evidence that the employment effects of JCS are likely to be negative which might be due to a missing market ability of JCS, possible stigma effects or locking-in effects.[9]

Looking at other wage subsidy programmes, e.g. SAM, SAM-East or EIM there are no microeconometric evaluation studies at all. The same holds for microeconometric studies trying to evaluate the employment effects of existing wage subsidy programs on the labour demand. Disney et al. (1992) use an aggregate framework analyzing the effects of various active labour market programmes on the unemployment flows for West Germany and find no significant effects of JCS and EIM. Schmid et al. (1999) using a similar methodological approach report positive effects regarding long-term unemployment. In contrast to studies trying to evaluate existing wage subsidy programmes there are at least some works aimed at assessing the impact of recent proposals for the reform of wage subsidies.[10] These studies use estimates of labour supply and demand elasticities in order to assess the effects of the various reform proposals on the employment and wage level.

---

[7]See e.g. Kölling (1998) or Nickel (1986).

[8]Hagen and Steiner (2000) and Steiner (2000) provide a good overview of the existing empirical studies.

[9]See Hujer and Caliendo (2000) or Hagen and Steiner (2000).

[10]See e.g. Schupp et al. (1999), Buslei and Steiner (1999) or Bender et al. (1999)

5. The IAB-Establishment Panel - Some Descriptive Results

In the following we will give a brief overview of the dataset used in our empirical analysis, namely the IAB establishment panel data set[11], and present some descriptive results about our sample.

The IAB establishment panel is a yearly survey of the demand side situation of the labour market. The unit of interest is the establishment, i.e. the local unit where the activity of a company takes place. The basis for the IAB establishment panel is the employment statistics register of the FEO, where employers have to report information regarding their employees subject to a social compulsory security scheme. Starting from this register a stratified representative sample is drawn. The IAB panel started in 1993 with about 4,200 establishments from West-Germany. Since 1996 establishments from East-Germany are also included in the dataset that in 1999 contained about 11,000 units.

The IAB panel is organized in modular form. There are topics covered annually like changes in the level and structure of employment and questions about the establishment policy, e.g. information about the business volume and investment. Other topics are only covered irregularily, e.g. information about innovations and also public employment subsidies. Additionally there is also a special questionnaire about actual questions included into the panel every year. Information about wage subsidies are provided for the years 1992, 1994, and 1996 to 1998.

Our evaluation strategy is a conditional difference-in-differences estimator, which will be explained in section 6. We will use information whether an establishment participated in a wage subsidy programme in 1996 or not, that is our treatment period is 1996 and our reference year is 1995. We then examine our outcome measure for three years, i.e. from 1997 to 1999. We will focus our attention therefore on those West German establishments which participated in all five waves from 1995 to 1999. The total number of these establishments is about 1,700 and we will refer to this group as our panel sample.

In order to analyze whether our panel sample is representative compared with the establishments which participated in every single wave, i.e. the cross-section establishments, we will compare them with respect to their affiliation to industrial sectors, their size class and the development of their number of employees.

If we compare the panel establishments with all establishments which participated e.g. in 1995 there are no major differences regarding the affiliation to industrial sectors and only small differences regarding the size of the firms. These differences increase somehow when comparing the establishments who participated from 1995 to 1999 with all establishments participating in 1999 especially with regard to the size class. Large firms are over-represented in the panel sample whereas smaller establishments with less than 50 employees are under-represented.

Looking at the development of the employment shows significant differences between the panel and the cross-section establishments (figure 3). Whereas establishments participating in all five waves from 1995 to 1999 were keeping their employment level almost stable, the employment level of establishments in the corresponding years decreased steadily from about 500 to under 300 employees on average.

The next step consists in comparing the participating establishments with the group of non-participating ones. Figure 4 reveals that there are considerable differences between the group

---

[11]See Kölling (2000) or Bellmann and Kölling (2000).

Figure 3: **Total Employment, Panel- vs. Cross-Section Establishments, 1995-1999**
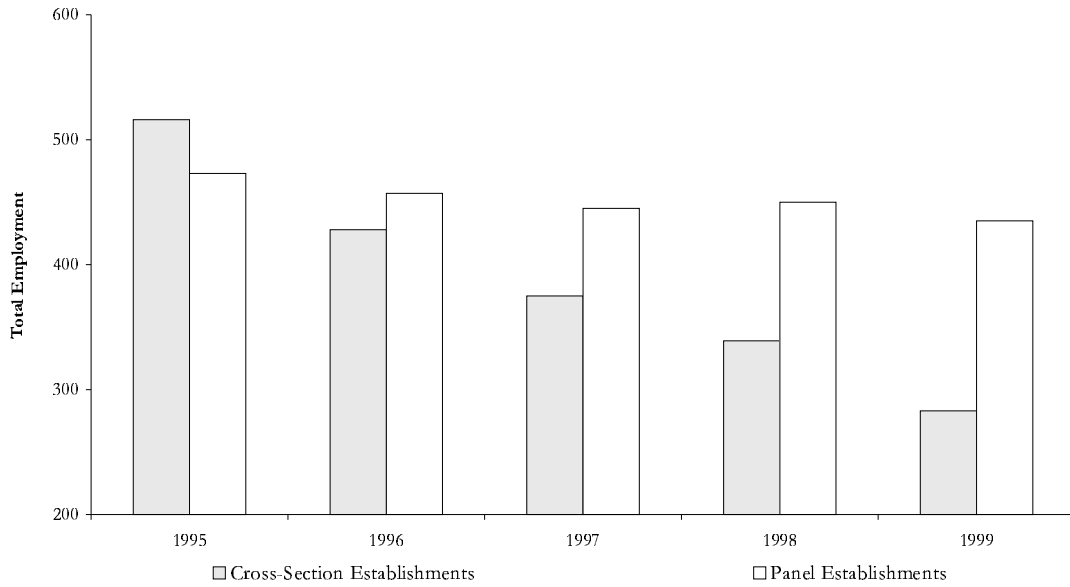
Figure 4: **Affiliation to Industrial Sectors, Participating vs. Non-Participating Establishments, 1996**
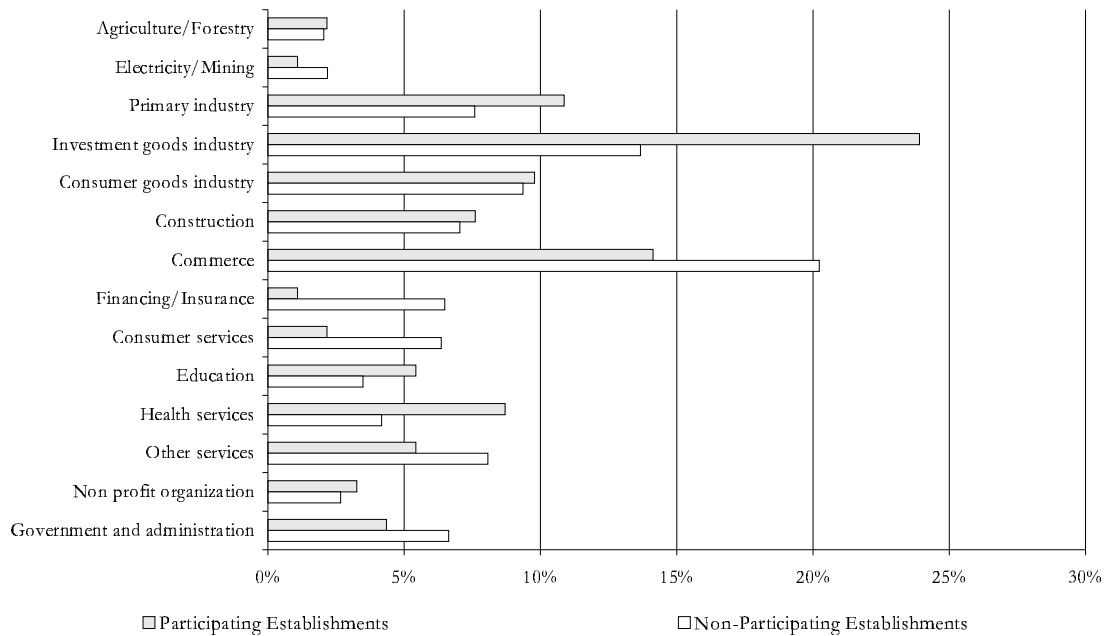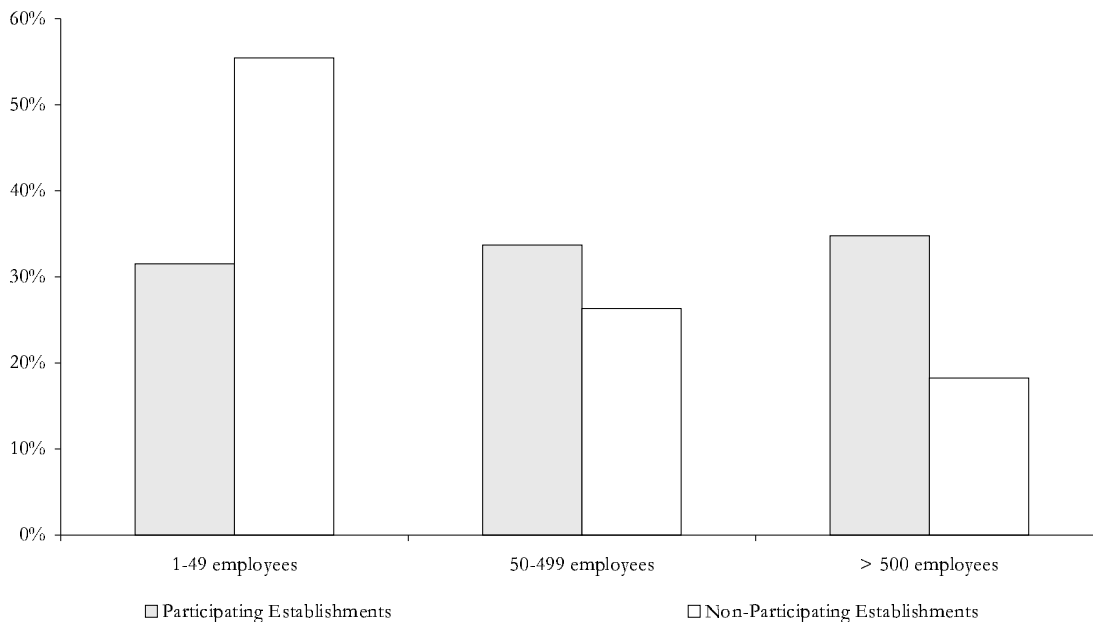
Figure 5: **Size Class, Participating vs. Non-Participating Establishments, 1996**

of participants and non-participants. When looking at the distribution of the establishments regarding the affiliation to industrial sectors one finds that especially establishments within the investment goods and primary industry are those asking for wage subsidies compared with the group of non-participants. On the other hand sectors like Commerce, Financing/Insurance or Consumer Services are under-represented within the treatment group.

These differences between participants and non-participants carry over to the size distribution (figure 5). Whereas establishments with less than 50, 50 to 499 and more than 500 employees are equally distributed in the group of treatment establishments, in the group of non participants small establishments are over-represented (55%), whereas huge establishments are under-represented (18%).

Therefore a simple comparison between establishments participating in a wage subsidy programme and non-participating establishments bears the risk of selection effects. Our matching approach tries to account for this kind of selectivity by constructing an adequate control group, i.e. comparing only comparable establishments.

6. The Evaluation Process

6.1. Outcome Measure

The first step and one major element of programme evaluation is the suitable choice of an outcome measure that is clarifying what outcomes are to be considered, i.e. what should be defined as a success. If we look at the labour supply side possible outcomes might be the earnings of an individual after training participation, or his employment situation. Schmidt (1999) notes that

10

there are some problems arising with the choice of an appropriate outcome measure as outcomes may not be comparable across interventions. Therefore a policy maker who has to decide which measure to implement will normally try to translate the gains of a programme into money terms or to carry out a so called cost-utility analysis. Still this is not easy because new problems like time or group preferences emerge.

As we are dealing not with individual persons but individual firms in our study, one obvious outcome measure is the labour demand of the firm. The closest approximation we get for this figure are the actual employed workers in the firm. As we are not interested in short-term effects of the wage subsidy we have to examine our outcome measure in a longer perspective. Therefore we compare the development in the employment statistics between participating and non-participating firms for three years after the programme took place. Beside the total employment we are also trying to estimate the effects on different heterogeneous skill groups. A conditional difference-in-differences approach will be used which we are going to describe in section 6.4.

But before we do so, we have to find a solution for the fundamental evaluation problem which arises from the fact that participants in active labour market programmes are a selected group and thus not directly comparable to the non-participants.

6.2. The Microeconometric Evaluation Problem

The framework serving as a guideline for our empirical analysis is the 'potential outcome approach', which is most often just called the Roy-Rubin-model (Roy (1951), Rubin (1974)). Inference about the impact of a treatment on the outcome of an individual involves speculation about how this individual would have responded, had he or she not received the treatment. In the basic model there are two potential outcomes (Y^T, Y^C) for each individual, where Y^T indicates a situation with training and Y^C without. This setup can be easily transferred to our context where the units under consideration are not individuals but establishments, from now on referred to as individual units. We define a binary assignment indicator D, indicating whether an individual unit actually participated in a programme $(D = 1)$ or not $(D = 0)$ (Hujer and Wellner (2000), Lechner (2000)). The treatment effect for each individual unit is then defined as the difference between its potential outcomes:

$$\Delta = Y^T - Y^C. \tag{7}$$

The fundamental problem of evaluating this individual treatment effect arises because the observed outcome for each individual unit is given by:

$$Y = D \cdot Y^T + (1 - D) \cdot Y^C. \tag{8}$$

Unfortunately we can never observe Y^T and Y^C for the same individual unit simultaneously. The unobservable component in (8) is called the counterfactual outcome, so that for units who participated in the measure $(D = 1)$, Y^C is the counterfactual outcome, and for those who did not it is Y^T.

Note that there will be no opportunity to ever estimate individual gains with confidence. Therefore we have to concentrate on the population average of gains from treatment. The most prominent evaluation parameter is the so-called average treatment effect on the treated:

11

$$E(\Delta \mid D = 1) = E(Y^T \mid D = 1) - E(Y^C \mid D = 1). \tag{9}$$

Like Hujer and Wellner (2000) note, this parameter gives an answer to the following question: "What is the expected, or mean outcome gain to individuals who received treatment to the hypothetical situation had they not received it?" This question focuses directly on actual participating units, so that it determines the realized gross gain from the programme and can be compared with its costs. This will help to decide whether the programme is a success or not (Heckman, Ichimura, Todd (1997, 1998), Heckman, LaLonde, Smith (1999)).

The second term on the right side in equation (9) is unobservable as it describes the hypothetical outcome without treatment for those units who received treatment. If the condition

$$E(Y^C \mid D = 1) = E(Y^C \mid D = 0) \tag{10}$$

holds, we can use the non-participants as an adequate control group. This identifying assumption is likely to hold only in social experiments, where the key concept is the randomized assignment of individual units into treatment and control groups. In nonexperimental data, equation (10) will normally not hold:

$$E(Y^C \mid D = 1) \neq E(Y^C \mid D = 0). \tag{11}$$

The use of the non-participants as a control group will therefore lead to a selection bias. Heckman and Hotz (1989) point out that selection might occur on observables or unobservables. We will present in the following sub-sections an estimation approach that tries to estimate the unobserved counterfactual term using the observed outcome information of the non-participants and taking into account selection on observables as well as selection on unobservables (see Hujer and Caliendo (2000) for a detailed description). We will start with the matching approach before we will present a conditional difference-in-differences estimator, suggested by Heckman, LaLonde, and Smith (1999) and recently applied on the labour supply side by Bergemann, Fitzenberger, Schultz, and Speckesser (2000).

It is worth noting that despite the fact that most evaluation research focuses on average outcomes, maybe because most statistical techniques focus on mean effects, there is also a growing interest regarding effects of policy variables on distributional outcomes. Examples where distributional consequences matter on the labour supply side include subsidized training programmes (LaLonde (1995)) or minimum wages (DiNardo, Fortin, and Lemieux (1996)). Koenker and Bilias (2000) recently showed that quantile regression methods can play a constructive role in the analysis of duration (survival) data, too. Clearly, if the data is rich enough it is worth to examine distributional effects for sub-groups of the targeted population on the labour demand side, too.

6.3. Selection on Observables and Unobservables

Matching is in one of the most appealing non-experimental approaches to solve the fundamental evaluation problem. That is because it shows a close link to the experimental context. The basic idea underlying the matching approach is to search from a large group of non-participants those individual units who are similar to the participants in all relevant pre-treatment characteristics. That being done, the differences in the outcomes between the well selected and thus adequate

control group and the participants can then be attributed to the programme.[12]. Following Rubin (1977) treatment assignment may be random given a set of covariates. The construction of a valid control group via matching is based on the identifying assumption that conditional on all relevant pre-training covariates (Z), the potential outcomes (Y^T, Y^C) are independent of the assignment to training. This so called conditional independence assumption (CIA) can be written formally as:

$$Y^T, Y^C \amalg D \mid Z.^{[13]} \tag{12}$$

If assumption (12) is fulfilled we get:

$$E(Y^C \mid Z, D = 1) = E(Y^C \mid Z, D = 0) = E(Y^C \mid Z) \tag{13}$$

Similar to randomization in a classical experiment, the role of matching is to balance the distributions of all relevant pre-treatment characteristics Z in the treatment and control group, and thus to achieve independence between the potential outcomes and the assignment to treatment, resulting in an unbiased estimate.[14] The implementation of conditioning on all relevant covariates is, however, limited in case of a high dimensional vector Z. For instance, if Z contains n covariates which are all dichotomous, the number of possible matches will be 2^n. In this case cell matching, that is exact matching on Z, is mostly not possible, because an increase in the number of variables increases the number of matching cells exponentially (Hujer and Wellner (2000)). To deal with this dimensionality problem, Rosenbaum and Rubin (1983) suggest the use of balancing scores $b(Z)$, i.e. functions of the relevant observed covariates Z such that the conditional distribution of Z given $b(Z)$ is independent of the assignment to treatment, that is $Z \amalg D \mid b(Z)$ holds.

For participants and non-participants with the same balancing score, the distributions of the covariates Z are the same, i.e. they are balanced across the groups. Moreover Rosenbaum and Rubin (1983) show that if the treatment assignment is strongly ignorable when Z is given, it is also strongly ignorable given any balancing score.[15] The propensity score, i.e. the probability of participating in a programme is one possible balancing score. It summarizes the information of the observed covariates Z into a single index function. Rosenbaum and Rubin (1983) show how the conditional independence assumption extends to the use of the propensity score so that

$$Y^C \amalg D \mid P(Z). \tag{14}$$

Therefore we get:

$$E(Y^C \mid P(Z), D = 1) = E(Y^C \mid P(Z), D = 0) = E(Y^C \mid P(Z)), \tag{15}$$

which allows us to rewrite the crucial term in the average treatment effect (9) as:

[12] The matching approach originated in the statistical literature (see Rubin (1974), (1977), (1979), Rosenbaum and Rubin (1983), (1985a), (1985b) or Lechner (1998))

[13] For the purpose of estimating the mean effect of treatment on the treated the assumption of conditional independence of Y^C is sufficient, because we like to infer estimates of Y^C for units with $D = 1$ from data on units with $D = 0$ (Heckman, Ichimura, and Todd (1997)).

[14] If we say relevant we mean all those covariates that influence the assignment to treatment as well as the potential outcomes.

[15] Strongly ignorable means that assumption (12) holds and: $0 < p(Z) \equiv P(D = 1 \mid Z) < 1$. The latter ensures that there are no characteristics in Z for which the propensity score is zero or one. Proofs can be found in Rosenbaum and Rubin (1983).

$$E(Y^C \mid D = 1) = E_{P(Z)}[E(Y^C \mid P(Z), D = 0) \mid D = 1]. \tag{16}$$

Hujer and Wellner (2000) note that the outer expectation is taken over the distribution of the propensity score in the treated population. The major advantage of the identifying assumption (14) is that it turns the estimation problem into a much easier task since one only has to condition on a univariate scale, i.e. on the propensity score. When $P(Z)$ is known the problem of dimensionality can be eliminated. The evaluation of the counterfactual term via matching on the basis of the group of non-participants then only requires us to pair participants with non-participants which have the same propensity score. This insures a balanced distribution of Z across both groups. Unfortunately $P(Z)$ will not be known a priori so it has to be replaced by an estimate, e.g. coming from a Probit model. Clearly the empirical power of matching to reduce the problem of selection bias relies crucially on the quality of the estimate of the propensity score on the one hand and on the existence of comparison units that have equal propensity scores as the treated units. If the latter is not ensured we face the risk of incomplete matching with biased estimates.

Recently it has been claimed that controlling for selection on observables may not be sufficient since remaining unobservable differences might still lead to a biased estimation of treatment effects.

These differences can be explained straightforward in the individual context: Participants might expect different benefits from participation in a treatment that might influence their decision to participate. Furthermore some groups might exhibit bad labour market prospects or differences in motivation. These things are unobservable to a researcher and might cause a selection bias.

In our context there might be such unobservable factors which influence the participation decision and/or the performance of the establishment, too. The quality of the management might be one such factor. Plans for future projects which are of course unobservable in nature or the competition structure of the market as well as certain economic conditions are other examples which show the importance of unobservable characteristics.

To account for selection on unobservables, Heckman, LaLonde, and Smith (1999) suggest econometric selection models and difference-in-differences (DiD) estimators. The DiD-estimator requires access to longitudinal data and can be seen as an extension to the classical before-after estimator (BAE). Whereas the BAE compares the outcomes of participants after they participate in the programme with their outcomes before they participate, the DiD-estimator eliminates common time trends by substracting the before-after change in non-participant outcomes from the before-after change for participant outcomes. The DiD-estimator is based on the assumption of time-invariant linear selection effects. The critical identifying assumption of this method is, that conditional on individual characteristics Z, the biases are the same on average in different time periods before and after the period of participation in the programme, so that differencing the differences between participants and non-participants eliminates the bias (Heckman, Ichimura, Smith, and Todd (1998)). Let t be a post-programme period and t' a pre-programme period. The expected outcome for an individual unit i at time t can be written as:

$$Y_{it} = \alpha_{it} + D_{it} \cdot Y_{it}^T + (1 - D_{it}) \cdot Y_{it}^C, \tag{17}$$

where α_{it} captures the effects of selection on unobservables. The DiD-estimator contrasts the

change for the participants i with the change for non-participants j:

$$[Y_{it}^T - Y_{it'}^C] - [Y_{jt}^C - Y_{jt'}^C]. \tag{18}$$

Its validity relies crucially on the assumption:

$$\alpha_{it} = \alpha_{it'}. \tag{19}$$

Only if the selection effect is time-invariant it can be cancelled out and an unbiased estimate will be achieved.[16]

6.4. Conditional Difference-in-Differences Approach

Before we start our estimation procedure we like to resume our basic ideas. First of all we have outlined that the firms which participate in a wage subsidy programme might differ systematically from non-participating firms. We can say, that the participation decision is selective. This selection might occur on observable or unobservable characteristics. Therefore an evaluation approach should take these selection effects into account. We follow a two-step-procedure:

Step 1: On a first level we estimate propensity scores, taking into account all relevant observable characteristics and try to find for every participating firm a comparable non-participating firm. This should ensure that both groups do not differ systematically regarding observable characteristics.

Step 2: As a second step we use a DiD-approach to estimate the treatment effect for three periods after the programme took place. The estimator can be written as:

$$
\begin{aligned}
\Delta_{t+1} &= [Y_{t+1}^T - Y_{t-1}^T] - [Y_{t+1}^C - Y_{t-1}^C] \\
\Delta_{t+2} &= [Y_{t+2}^T - Y_{t-1}^T] - [Y_{t+2}^C - Y_{t-1}^C]. \\
\Delta_{t+3} &= [Y_{t+3}^T - Y_{t-1}^T] - [Y_{t+3}^C - Y_{t-1}^C]
\end{aligned}
\tag{20}
$$

The conditional DiD-approach should ensure, that Δ estimates the true treatment effect, as we have controlled for all observable and unobservable characteristics. Figure 7 in the appendix tries to illustrate the idea for a very simple case.

Let us assume that we have an establishment (firm 1) that participated in a programme at time t. With our matching approach we have been able to find a comparable establishment (firm 2) out of the groups of non-participants which is very similar to firm 1 regarding the total employment, but did not participate at all in such a programme.[17] We assume that there has been a recession in period t that has lead to (the same) negative employment effect for all establishments. Beside that, there are no other influences affecting the total employment situation, that is the total employment remains constant over the next years. Firm 1 participated in a programme in t. The programme effect outnumbers the effects of the recession and leads to a higher employment level compared to period $t - 1$. If we concentrate on firm 1 and make a simple before-after comparison between

---

[16]The differencing leads to: $Y_{it} - Y_{it'} = [D_{it} \cdot Y_{it}^T + (1 - D_{it}) \cdot Y_{it}^C] - [D_{it'} \cdot Y_{it'}^T + (1 - D_{it'}) \cdot Y_{it'}^C] + [\alpha_{it} - \alpha_{it'}]$. If $\alpha_{it} = \alpha_{it'}$ the last term in the expression is cancelled out.
[17]For the sake of simplicity we think of total employment as our only matching variable.

$t+1$ and $t-1$ we would conclude, that the programme did not have any effect at all. But if we look instead at the DiD-Estimator we see, that it is capable to evaluate the true treatment effect correctly, because it takes into account the drop in employment caused by the recession in t (through the comparison with the development of firm 2).

6.5. Estimating the Propensity Score

In the following we will estimate a Probit model to obtain the probabilities for the various establishments to participate at a wage subsidy programme in 1996. Unfortunately we could not take into account heterogeneous treatments, because the number of observations for every treatment has been to small. Therefore we had to pool all employment integration measures, including the programmes for long-term and elderly unemployed, with job creation schemes. This pooling leaves us with a treatment group of 87 establishments. In addition we have 1,354 establishments which did not participate at all in a programme and were surveyed between 1995 and 1999.[18]

The Probit model can be written as:

$$y_i^* = \beta' z_i + u_i, \tag{21}$$

with:

$$y_i = \left\{ \begin{array}{ll} 1, & \text{if } y_i^* > 0 \\ 0, & \text{otherwise.} \end{array} \right. \tag{22}$$

y_i^* is a latent variable defined as a function of a set of relevant participation determinants, which are captured by the vector z_i, whereas the observable y_i equals one if the i-th establishment participates in one of the above listed programmes and zero otherwise.

It is assumed that u_i is normally distributed, i.e. $u_i \sim N(0, \sigma^2)$ and that u_i is independent of the z_i. The probability for $y_i = 1$, P_i, conditional on z_i is therefore:

$$
\begin{aligned}
P(y_i = 1 \mid z_i) &= P\left(\frac{u_i}{\sigma} > \frac{-\beta' z_i}{\sigma} \mid z_i\right) \\
&= \Phi\left(\frac{\beta' z_i}{\sigma} \mid z_i\right). \tag{23}
\end{aligned}
$$

Where Φ is the cumulative distribution function of the standard normal distribution. One can see that the expression $\frac{\beta'}{\sigma}$ is not identified so that in the following we will set σ equal to one.

The standard error of the predicted probability is calculated using the asymptotic covariance matrix V as follows:

$$Var(P_i) = \phi(\beta' z_i)^2 z_i V z_i. \tag{24}$$

The estimation results can be found in Table 3 in the appendix. Beside the 11 dummy variables for the industrial sectors we found only one further variable to be statistically significant on a

---

[18]One should note that we have only 1,441 establishments left for our analysis. Compared to the number of establishments which have been surveyed between 1995 and 1999 (about 1,700) the loss is caused by missing values for the outcome variables.

10%-level, namely the number of other employees.[19] The dummy variable for industrial sector 8 (Financing / Insurance) was used as the reference category. The so obtained propensity scores together with the standard errors will be used in the following matching approach.

6.6. The Matching Procedure

The matching procedure intends to find for every unit in the treatment group a comparable unit out of the control group. Differences in both groups regarding the observable characteristics should be minimized. The used algorithm for our matching procedure can be found in Table 4 in the appendix and is based on Rosenbaum and Rubin (1985b).

The empirical power of matching to reduce the problem of selection bias relies, as we have already noted, crucially on the quality of the estimate of the propensity score on the one hand and on the existence of comparison units that have equal propensity score as the treated units.

As we have only found a few parameters in our Probit model to be significant, we might have to impose further matching conditions, to ensure a good matching quality. In our comparison between participating and non-participating establishments in section 5 we have found that there are differences between the groups regarding the total employment as well as the affiliation to industrial sectors. Therefore we are conditioning in a second matching approach not only on the propensity score but also on the size of the establishment and the industrial sector.

Table 2 compares some characteristics between the participants (i.e. our treatment group) and the matched non-participants (i.e. our control group) as well as the group of all non-participants in the relevant year. In the first column we find the mean values of total employment, employment in the different skill groups, the trade volume and the wage costs for the treatment group (87 units).

Table 5: **Comparison Between Participants and Matched Non-Participants**[a]

	Partici- pants	Match-1: Score	Match-2: Score, ES and IS	All Non- Partici- pants
No. of individual units	87	86	77	1354
Variable	**Mean**	**Mean**	**Mean**	**Mean**
Total Employment	576.4	448.7	578.4	421.2
High-Skilled Employees	381.9	290.7	370.9	291.9
Low-Skilled Employees	168.3	137.1	181.4	112.2
Other Employees	26.2	20.9	26.1	17.1
Wage Costs (in thousand DM)	2558.0	2232.4	2778.9	2053.6
Trade Volume (in million DM)	331.5	289.2	645.5	901.5

[a] ES stands for the size of the establishment regarding total employment (three categories: 1-49, 50-499, over 500), whereas IS indicates the industrial sector.

---

[19]Note that we did not include the dummy variables for the industrial sectors 2 (Electricity / Mining) and 9 (Consumer Services) in our estimation, because no firm in the treatment group has been in either one of these industrial sectors.

If we compare these results with column 4, that is the mean values for all non-participants we can see that there are great differences, regarding not only the employment, but also the trade volume and the wage costs. In the second column (Match-1) we have the results for our first matching approach, namely the matching (only) on the propensity score. We see that the matched control group is more similar to the treatment group, not only regarding the employment statistics but also regarding the trade volume. But still the matching has not been completely successful, as there are still some differences regarding e.g. the low-skilled employees. Turning to column 3 and our second matching approach (Match-2), that conditions in addition to the propensity score on the size of the establishment and the industrial sector, we find the mean values in nearly all employment groups to be very similar to the means in the treatment group. The matching quality is improved and we will use these matched pairs for our DiD-estimation.[20]

6.7. Results of the Conditional DiD-Estimation

The results of our DiD-estimation for the different skill groups are depicted in figure 6. As the most actual wave of the IAB establishment panel is not available for research yet, we could examine the effects for three years until 1999. Leaving the group of other employees aside, we can see that the development in both skill groups is negative.

Figure 6: **The DiD-Estimator for Different Employment Groups, 1997-1999**
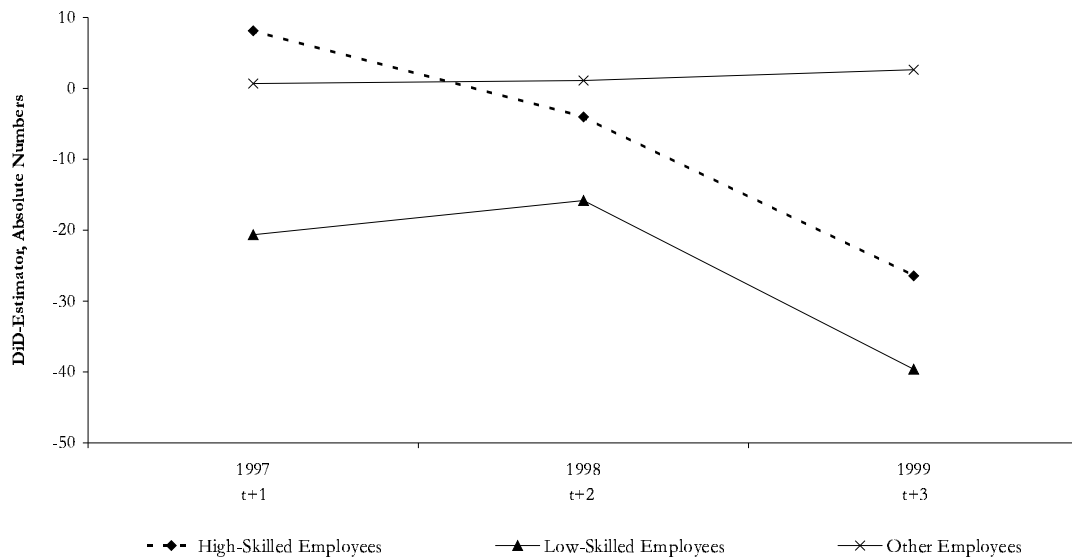
The development for the group of high-skilled workers is positive in the first year, but turns negative in the two following years. As we have compared the situation of participating and non-participating establishments before and after the programme took place this means the following: The development from 1995 to 1999 in the employment for high-skilled workers has been worse for participating establishments compared to non-participating establishments, by the absolute number of 26 workers.

---

[20]One should note that the price for the better matching quality is a loss in observations. Whereas we could find 86 suitable controls in our first approach, this is the case only for 77 establishments now.

The situation for the low-skilled workers is similar, with the difference, that the DiD-estimator is negative throughout the whole period. In 1997 and 1998 the difference lies around 20 workers, before it rises to nearly 40 workers in the third year.

The results do not show any positive long-term effects on the employment situation of high- or low-skilled workers. As the wage subsidies are generally aimed at workers with bad labour market prospects, one might have expected at least a positive effect for low-skilled workers, but that could not be found. This might however be due to the fact, that the wage subsidies are aimed only at a subgroup of low-skilled workers, e.g. long-term unemployed, and therefore displacement effects might have occurred. However, the results have to be treated with caution. First of all, we have not been able to take into account heterogeneous treatments. Even though all programmes have some common goals, the effects of each programme might differ. Furthermore the pooling of programmes makes it difficult to assess the duration of every treatment for every particular establishment, as the possible duration ranges between 6 and 36 months. In addition the limited number of periods makes it difficult to find out the long-term effects, because some programmes might have just finished. A further problem is the small group of participants, making it difficult to generalize the results.

7. Conclusion

We were interested in estimating the effects of wage subsidy programmes on the labour demand. After a short outline of existing programmes in Germany we showed, that the theoretical effects of these programmes are ambiguous and an empirical evaluation is needed. In section 5 we presented the dataset used for our analysis, namely the IAB establishment panel. A first descriptive analysis showed, that participating and non-participating establishments differ systematically. This well known selection problem in microeconometric evaluation studies might be caused by observable or unobservable characteristics. Therefore we used a conditional difference-in-differences approach to account for both. The results do not show any positive employment effects either on the high- or low- skilled workers. Out of several reasons, like the limited time period and the small number of observations in our treatment group, the results have to be treated with caution.

One major task for the future is the accounting for heterogeneous treatments, what might be possible with the following waves and the extension of the evaluation to East Germany.

References

BELLMANN, L., AND A. KÖLLING (2000): "The IAB-Establishment Panel as an Instrument for Scientific Research," Paper prepared for the Ninth International Conference on Panel Data, June 22th - 23th, Geneva, Switzerland.

BENDER, S., B. KALTENBORN, H. RUDOLPH, AND U. WALWEI (1999): "Förderung eines Niedriglohnsektors - Die Diagnose stimmt, die Therapie noch nicht," *IAB-Kurzbericht*, 6.

BERGEMANN, A., B. FITZENBERGER, B. SCHULTZ, AND S. SPECKESSER (2000): "Multiple Active Labor Market Policy Participation in East Germany: An Assesment of Outcomes," Working Paper, Institute for Economic Research Halle, University of Mannheim.

BUNDESANSTALT FÜR ARBEIT (2000): *ANBA - Daten zu den Eingliederungsbilanzen 1999*. Nürnberg.

BUSLEI, H., AND V. STEINER (1999): *Beschäftigungseffekte von Lohnsubventionen im Niedriglohnbereich*. Nomos Verlagsgesellschaft, Baden-Baden.

DINARDO, J., N. FORTIN, AND T. LEMIEUX (1996): "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach," *Econometrica*, 64, 1001–1045.

DISNEY, R., L. BELLMANN, A. CARRUTH, W. FRANZ, J. JACKMAN, R. LAYARD, H. LEHMANN, AND J. PHILPOTT (1992): *Helping the Unemployed*. London.

EICHLER, M., AND M. LECHNER (1999): "An Evaluation of Public Employment Programmes in the East German State of Sachsen-Anhalt," Discussion Paper No.94, IZA, Bonn.

FITZENBERGER, B., AND S. SPECKESSER (1998): "Zur wissenschaftlichen Evaluation der Aktiven Arbeitsmarktpolitik in Deutschland," *Mitteilungen aus der Arbeitsmarkt und Berufsforschung, Schwerpunktheft: Erfolgskontrolle aktiver Arbeitsmarktpolitik*, 3, 532–549.

HAGEN, T., AND V. STEINER (2000): *Von der Finanzierung der Arbeitslosigkeit zur Förderung von Arbeit - Analysen und Empfehlungen zur Arbeitsmarktpolitik in Deutschland*. Nomos Verlagsgesellschaft, Baden-Baden.

HAMERMESH, D. S. (1993): *Labor demand*. Princeton University Press, Princeton.

HECKMAN, J., AND J. HOTZ (1989): "Choosing Among Alternative Nonexperimental Methods for Estimating the Impact of Social Programs: The Case of Manpower Training," *Journal of the American Statistical Association*, 84, 862–880.

HECKMAN, J., H. ICHIMURA, J. SMITH, AND P. TODD (1998): "Characterizing Selection Bias Using Experimental Data," *Econometrica*, 66, 1017–1098.

HECKMAN, J., H. ICHIMURA, AND P. TODD (1997): "Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme," *Review of Economic Studies*, 64, 605–654.

————— (1998): "Matching as an Econometric Evaluation Estimator," *Review of Economic Studies*, 65, 261–294.

HECKMAN, J., R. LALONDE, AND J. SMITH (1999): "The Economics and Econometrics of Active Labor Market Programs," in *Handbook of Labor Economics Vol.III*, ed. by O. Ashenfelter, and D. Card, pp. 1865–2097. Elsevier, Amsterdam.

HÜBLER, O. (1997): "Evaluation beschäftigungspolitischer Maßnahmen in Ostdeutschland," *Jahrbücher für Nationalökonomie und Statistik*, 216, 21–44.

HUJER, R., AND M. CALIENDO (2000): "Evaluation of Active Labour Market Policy: Methodological Concepts and Empirical Estimates," Discussion Paper No.236, IZA, Bonn.

HUJER, R., K. MAURER, AND M. WELLNER (1999): "The Effects of Public Sector Sponsored Training on Unemployment Duration in West Germany - A Discrete Hazard Rate Model Based on a Matched Sample," *ifo-Studien*, 3/1999, 45 Jhg., 371–410.

HUJER, R., AND M. WELLNER (2000): "The Effects of Public Sector Sponsored Training on Individual Employment Performance in East Germany," Discussion Paper No.141, IZA.

KOENKER, R., AND Y. BILIAS (2000): "Quantile Regression For Duration Data:A Reappraisal Of The Pennsylvania Reemployment Bonus Experiments," Working Paper, University of Illinois.

KÖLLING, A. (1998): *Anpassungen auf dem Arbeitsmarkt. Eine Analyse der dynamischen Arbeitsnachfrage in der Bundesrepublik Deutschland.* Institut für Arbeitsmarkt- und Berufsforschung, Beiträge zur Arbeitsmarkt- und Berufsforschung, Nürnberg.

——— (2000): "The IAB-Establishment Panel," *Schmollers Jahrbuch*, 120, 291–300.

KRAUS, F., P. PUHANI, AND V. STEINER (1999): "Employment Effects of Publicly Financed Training Programs - The East German Experience," *Jahrbücher für Nationalökonomie und Statistik*, 219, 216–248.

KRAUS, F., AND V. STEINER (1995): "Haben Teilnehmer an Arbeitsbeschaffungsmaßnahmen in Ostdeutschland bessere Wiederbeschäftigungschancen als Arbeitslose?," in *Mikroökonomik des Arbeitsmarktes*, ed. by L. Bellmann, and V. Steiner, pp. 387–423. IAB - Beiträge zur Arbeitsmarkt- und Berufsforschung 192.

LALONDE, R. (1995): "The Promise of Public-Sector Sponsored Training Programs," *Journal of Economic Perspectives*, 9, 149–168.

LECHNER, M. (1998): "Mikroökonometrische Evaluationsstuidien: Anmerkungen zu Theorie und Praxis," in *Qualifikation, Weiterbildung und Arbeitsmarkterfolg. ZEW-Wirtschaftsanalysen Band 31*, ed. by F. Pfeiffer, and W. Pohlmeier. Nomos-Verlag.

——— (2000): "An Evaluation of Public Sector Sponsored Continuous Vocational Training Programs in East Germany," *Journal of Human Resources*, Spring, 347–375.

NICKEL, S. J. (1986): "Dynamic Models of Labour Demand," in *Handbook of Labor Economics, Volume I*, ed. by O. Ashenfelter, and R. Layard, pp. 473–522. North-Holland.

ROSENBAUM, P., AND D. RUBIN (1983): "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrica*, 70, 41–50.

——— (1985a): "The Bias due to incomplete Matching," *Bioemetrics*, 41, 103–116.

——— (1985b): "Constructing a Control Group Using Multivariate Matched Sampling Methods that Incorporate the Propensity Score," *The American Statistican*, 39, 33–38.

ROY, A. (1951): "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 3, 135–145.

RUBIN, D. (1974): "Estimating Causal Effects to Treatments in Randomised and Nonrandomised Studies," *Journal of Educational Psychology*, 66, 688–701.

——— (1977): "Assignment to Treatment Group on the Basis of a Covariate," *Journal of Educational Studies*, 2, 1–26.

——— (1979): "Using Multivariate Matched Sampling and Regression Adjustment to Control Bias in Observational Studies," *Journal of the American Statistical Association*, 74, 318–328.

SCHMID, G., H. MOSLEY, C. HILBERT, AND H. SCHÜTZ (1999): "Zur Effektivität aktiver Arbeitsmarktpolitik - Erfahrungen aus der integrierten Wirkungs- und Implementationsstudie," Discussion paper, MittAB, Nürnberg.

21

SCHMIDT, C. (1999): "Knowing What Works - The Case for Rigorous Program Evaluation," Discussion Paper No.77, IZA.

SCHUPP, J., J. VOLZ, G. WAGNER, AND R. ZWIENER (1999): "Zuschüsse zu den Sozialversicherungsbeiträgen im Niedriglohnbereich: wenig zielgerichtet und teuer," *DIW Wochenbericht*, 27/99.

STAAT, M. (1997): *Empirische Evaluation von Fortbildung und Umschulung - Schriftenreihe des ZEW 21*. Nomos Verlagsgesellschaft, Baden-Baden.

STEINER, V. (2000): "Employment Effects of Social Security Reforms in Germany," Habilitationsschrift.

A Appendix

Table 3: **Estimation Results of the Probit-Model for the Propensity Score**

Variable	Estimate	Std. Err.
Intercept	-2.6672543	0.37929*
Number of high-skilled employees	0.00005063	0.000124
Number of low-skilled employees	-0.0000271	0.000175
Number of other employees	0.00299062	0.001729*
Wage costs in DM	-2.214E-8	2.078E-8
Trade Volume in DM	-1.683E-12	1.65E-11
Number of employees dismissed	0.0006591	0.002178
Dummy variables, 1 if condition is fulfilled, 0 else		
Parts of the establishments closed	-5.3680712	8257.907
Parts of the establishments separated out	-0.3224929	0.407369
Parts of the establishment outsourced	0.36806154	0.579347
Number of employees expected to decline in the future	0.07626128	0.141302
Trade volume expected to decline in the future	0.03251574	0.207533
Investment expected to decline in the future	0.02071342	0.130548
Bad economic situation	-0.0446811	0.124427
Dummy Variables for the industry affiliation, 1 if establishment is in branch, 0 else.		
Agriculture/Forestry	1.15172269	0.513764*
Primary industry	1.28021957	0.416152*
Investment goods industry	1.4076258	0.391685*
Consumer goods industry	1.21006026	0.410126*
Construction	1.17664933	0.419922*
Commerce	0.99239678	0.395583*
Financing/Insurance	Reference	
Education	1.3569962	0.445622*
Health Services	1.39128327	0.428050*
Other Services	0.9567638	0.426293*
Non profit organization	0.09285134	0.036577*
Government and Administration	0.06233026	0.031336*
Log-Likelihood:	-307.19	

* denotes significance on a 10%-level.

Table 4: **The Matching Procedure**

Step 1:	Separate all units into two groups, with respect to the fact if they participated in a wage subsidy programme (treatment group) or not (control group).
Step 2:	Order the units in the treatment group randomly.
Step 3:	Take the first unit in the treatment group (denoted by i and estimate the propensity score $\hat{\beta}'z_i$ using the Probit-model specified in section 6.5. Estimate an interval $\hat{\beta}'z_i \pm c\sqrt{Var(\hat{\beta}'z_i)}$ and choose c accordingly, that a 90%-confidence interval around the score is achieved.
Step 4:	Find all units in the control group (denoted by j) that fulfill the condition: $\hat{\beta}'z_j \in \hat{\beta}'z_i \pm c\sqrt{Var(\hat{\beta}'z_i)}$
Step 5:	(a) If there is no unit in the control group that fulfills the condition in step 4, delete the unit in the treatment group and go back to step **3**. (b) If there is more than one unit that fulfills the condition estimate the difference $D(j,i) = |\hat{\beta}'z_j - \hat{\beta}'z_j|$ between the scores and take the unit out of the control group which has the smallest difference.
Step 6:	Take the two 'matched' units out of the groups and go back to Step **3**.

Figure 6: **The Conditional Difference-in-Differences Approach**